NTP's Modified One-Generation Reproduction Study

NTP Board of Scientific Counselors Meeting

April 13, 2011

Presenter: Dr. Paul Foster, Chief, NTP Toxicology Branch

Background

How the NTP evaluates reproductive toxicity – the multigeneration reproduction study.

The classical study employed to evaluate reproductive toxicity is the multigeneration reproduction study (e.g., Environmental Protection Agency Office of Prevention, Pesticides and Toxic Substances [EPA OPPTS] 870.3800, Organization for Economic Co-operation and Development [OECD] 416; Figure 1), where reproductive performance is evaluated over two breeding generations, typically in the rat. This is the only standard design where exposure to a test article occurs throughout the reproductive cycle (see Figure 2). The major objective of the study has been focused on fertility and fecundity of both parents and their offspring. Since the overall intent is to evaluate the ability of animals to mate and produce viable offspring, the study produces large numbers of animals for evaluation. Indeed, it is the largest and one of the most logistically complex regulatory-type studies routinely conducted to produce data for human health risk assessment. The NTP uses a variant of this basic design, termed the reproductive assessment by continuous breeding (RACB), which evaluates multiple pairings and litters in each generation (Chapin and Sloane, 1997).

The need for better evaluation of postnatal outcomes – improvements made to our multigeneration study.

As our knowledge of critical windows of exposure has increased, particularly with the increased focus on agents that may have endocrine-like activity, the last 15 years has shown the need for a study where there has been a larger focus on the evaluation of potential postnatal adverse outcomes. Thus, there have been updates to standard designs to incorporate more functional end points (e.g., sperm and oocyte analysis, vaginal cytology, indices of puberty and sexual differentiation) to improve the detection of agents affecting reproduction and the endocrine status of animals. In particular, in current study designs, the ability to evaluate (both detection and analysis of dose response) abnormalities of the reproductive tract routinely following *in utero* exposure to agents with endocrine activity was determined to be underpowered by several research groups (Blystone *et al.* 2010; Hotchkiss *et al.* 2008; McIntyre *et al.* 2002). For example, in an evaluation of prenatal developmental toxicity, every fetus is examined for potential abnormalities (typically ~ 250 fetuses per group) whereas in the multigeneration study, only one male and female pup per litter from a minimum of 20 litters is examined at adulthood for adverse pathological events (i.e., only 40 of the potential 250 animals/group produced). Some of NTP's own studies have shown the added value and increased statistical power of evaluating more offspring per litter by retaining them to adulthood, rather than discarding animals already produced, or performing only a gross examination at weaning, when the

reproductive organs are not fully differentiated or developed (Blystone *et al.* 2010). The NTP has already adopted the improved use of these animals, in its RACB study design, by carrying more animals through to adulthood for examination, rather than discarding them.

Movement to the inclusion of perinatal exposure periods in NTP rat cancer bioassays – need for a suitable preliminary study.

Following a workshop held by the NTP on the evaluation of tumors of the endocrine system (Thayer and Foster 2007), the NTP adopted a new default paradigm for its rat cancer bioassays that incorporates exposure during the perinatal period (i.e., gestation and lactation). The NTP has conducted a number of these "perinatal bioassays" in the past, but this would normally require a specific, scientific justification. Our current paradigm is that a perinatal study will be conducted, unless there is a specific, scientific reason why this should not occur. Before embarking on such a study, it would be customary to undertake a preliminary study that evaluated target organ toxicity (for a conventional cancer study, this would be the 90-day toxicity study) and enabled suitable dose levels to be selected for the cancer bioassay. Thus, for a long- term study involving exposure during pregnancy and lactation, a shorter duration study that involved exposure during these critical developmental windows would be required. The NTP has previously used a study that has exposure for 90 days after the pups are weaned. It became apparent that this basic design could easily be adapted to provide more detailed information on reproduction and development, as well as all the necessary information to select dose levels for the cancer study. Thus, the NTP would gain information on a greater array of toxicity end points and maximize the utility of the animals that had already been produced, such that the other, "stand alone" reproduction and developmental toxicity studies would not be required.

Other international efforts to develop new reproductive toxicity study designs.

In parallel with some of these NTP efforts, other initiatives had been taken internationally to refine reproductive toxicity testing for agrochemicals (Cooper *et al.* 2006) for which a large database of toxicity information is usually available. The study design (see Figure 3) published by Cooper et al had been proposed as a replacement for the EPA/OECD multigeneration reproduction study and in addition explored the incorporation of some additional end points (particularly developmental neurotoxicity and immunotoxicity) that would normally be conducted in separate, triggered assays. In Europe, the advent of REACH (Registration, Evaluation, Authorization and restriction of CHemical substances) was likely to require increased toxicity testing of chemicals at the same time that other efforts were seeking to reduce experimental animal usage. Significant attention was therefore focused on those assays that required (or produced) the largest number of experimental animals – the multigeneration reproduction study and the prenatal developmental toxicity study (usually conducted in two species). The OECD took up the challenge of finding a study that would provide adequate information for the evaluation of reproductive toxicity, but would also reduce animal numbers and has proposed a study (the extended one-generation reproduction study, based on the Cooper et al design) for adoption (see figure 4; the latest draft is available on line at *http://www.oecd.org/dataoecd/23/10/46466062.pdf*). The OECD design has undergone a

number of iterations and is in final draft form. It is proposed to be used for all chemicals (not just pesticides), where a much reduced toxicity database is likely to be available.

The proposed NTP design

The NTP design employs pregnant animals with dosing commencing at implantation (gestation day [GD] 6 in the rat) and continually exposes the dams throughout gestation and lactation. At weaning (usually postnatal day [PND] 21 in the rat), the offspring would be continued to be administered the test article at the same dose level as their respective dam and are subsequently assigned to a number of different cohorts that can be considered as interchangeable "cassettes" that can be included, or not, based on the NTP nomination, or other available information. These cassettes are essentially protocols used on other standard studies and would normally include:

- An evaluation of target organ toxicity, pathology, clinical pathology etc., similar to NTP's current 90-day toxicity protocol – **a subchronic toxicity cohort**. This would normally require 10 animals to be evaluated per sex, per dose group.

- An evaluation of prenatal developmental toxicity – **a teratology cohort**. One male and female offspring from each litter would be selected and non-sibling matings would be performed in each group on reaching sexual maturity (~PND 100). Just prior to expected delivery, a Cesarean section would be performed on the pregnant dams for a standard evaluation of external, visceral and skeletal abnormalities of the fetuses.

- An evaluation of breeding performance – **a littering cohort**. One male and female offspring from each litter would be selected and non-sibling matings would be performed in each group on reaching sexual maturity (~PND 100). The pregnant dams would be allowed to deliver their litters and raise them to weaning.

On a specific case basis, other cassettes could be added, or substituted, in the protocol, including an assessment of developmental immunotoxicity and/or developmental neurotoxicity, using protocols from studies previously employed by the NTP.

The study would normally be conducted in the NTP's default rat strain (the Harlan Sprague-Dawley). A sufficient number of time-mated animals would be acquired to ensure a minimum of 20 litters per dose group with normally 3 dose levels, plus a vehicle control. The normal route of exposure employed on such studies would be oral (e.g., dosed feed, drinking water, or gavage) and treatment would be continuous throughout the study (for gavage, direct dosing of pups may be required at least from PND 12, or as appropriate from toxicokinetic [TK] information). If no other toxicity information is available, a pilot study with a small number of pregnant dams would be required to set dose levels and potentially acquire preliminary TK data in pregnancy and lactation.

This NTP design emphasizes a full evaluation of the F_1 animals in the study. This represents a unique exposure group compared to other toxicity studies (i.e., exposure from implantation until adulthood). The design uses significantly fewer animals than a RACB multigeneration

reproduction study and generates important information on both reproduction and postnatal development, together with a pathological evaluation of all the offspring (after PND 4) when they reach adulthood to improve our ability to detect postnatal effects. A major addition will be the information achieved on prenatal developmental toxicity. The teratology and littering cohorts of animals will also allow the evaluation of fertility and fecundity and importantly, we can maintain the relationship between structural changes in the reproductive organs and functional outcomes in the same animals. In addition, the NTP will be able to maintain a 10-week exposure period prior to mating in the F_1 animals to ensure that any potential male germ cell effects could be reflected in a functional outcome.

Advantages of the NTP proposal over the draft OECD design

The OECD extended one generation study (Figure 4), while using fewer animals than a typical multigeneration study, does have a number of significant design flaws that have been highlighted at various recent scientific meetings (e.g., the Teratology Society/European Teratology Society exchange lectures in 2010). Some of these issues include: 1) a truncated exposure period before breeding the F_0 animals that will only produce functional effects in males if the agent under investigation affects epididymal sperm; 2) a failure to routinely breed F_1 animals and so not provide a full characterization of this unique exposure group; 3) the employment of internal triggers to make decisions on when breeding of the F_1 generation, or utilization of specific cohorts of animals, will occur. In particular, the triggers to be employed are not well described and do not take into account the timing for the production of data from the various components of the study, such that information will be available to make informed decisions on whether to apply triggers, or not; 4) the adverse structural changes will not be related to functional outcomes in the same animals if the triggers are not applied correctly; and 5) the neurotoxicity cohort is underpowered with only 10 animals/sex/group evaluated. Using this proposed design, one could conduct a reproduction study that does not adequately evaluate fertility and fecundity in any generation.

Conclusions

The proposed NTP study design represents an evaluation of reproductive and developmental toxicity while maximizing the utility of the animals available for study and reducing the overall number of animals produced compared to a multigeneration study. The number of animals employed is comparable to the recently proposed OECD design, but provides much more information on developmental outcomes. In addition, the design will facilitate NTP's requirement for information on subchronic toxicity and dose-setting before embarking on a perinatal chronic toxicity and carcinogenesis assay in the rat. Thus, the NTP will be able to refine our toxicity study designs, reduce overall animal use, and replace certain other standard toxicity assays by folding them into this proposed protocol.

References

Blystone, C. R., Kissling, G. E., Bishop, J. B., Chapin, R. E., Wolfe, G. W., and Foster, P. M. (2010). Determination of the di-(2-ethylhexyl) phthalate (DEHP) NOAEL for reproductive

development in the rat: Importance of the retention of extra animals to adulthood. *Toxicol Sci* **116**, 640-646.

Chapin, R. E., and Sloane, R. A. (1997). Reproductive assessment by continuous breeding: evolving study design and summaries of ninety studies. *Environ Health Perspect* **105 Suppl 1**, 199-205.

Cooper, R. L., Lamb, J. C., Barlow, S. M., Bentley, K., Brady, A. M., Doerrer, N. G., Eisenbrandt, D. L., Fenner-Crisp, P. A., Hines, R. N., Irvine, L. F., Kimmel, C. A., Koeter, H., Li, A. A., Makris, S. L., Sheets, L. P., Speijers, G., and Whitby, K. E. (2006). A tiered approach to life stages testing for agricultural chemical safety assessment. *Crit Rev Toxicol* **36**, 69-98.

Hotchkiss, A. K., Rider, C. V., Blystone, C. R., Wilson, V. S., Hartig, P. C., Ankley, G. T., Foster, P. M., Gray, C. L., and Gray, L. E. (2008). Fifteen years after "Wingspread"--environmental endocrine disrupters and human and wildlife health: where we are today and where we need to go. *Toxicol Sci* **105**, 235-59.

McIntyre, B. S., Barlow, N. J., and Foster, P. M. D. (2002). Male rats exposed to linuron in utero exhibit permanent changes in anogenital distance, nipple retention, and epididymal malformations that result in subsequent testicular atrophy. *Toxicol Sci* **65**, 62-70.

Thayer, K. A., and Foster, P. M. (2007). Workgroup report: National Toxicology Program workshop on Hormonally Induced Reproductive Tumors - Relevance of Rodent Bioassays. *Environ Health Perspect* **115**, 1351-6.

EPA Multigeneration Study

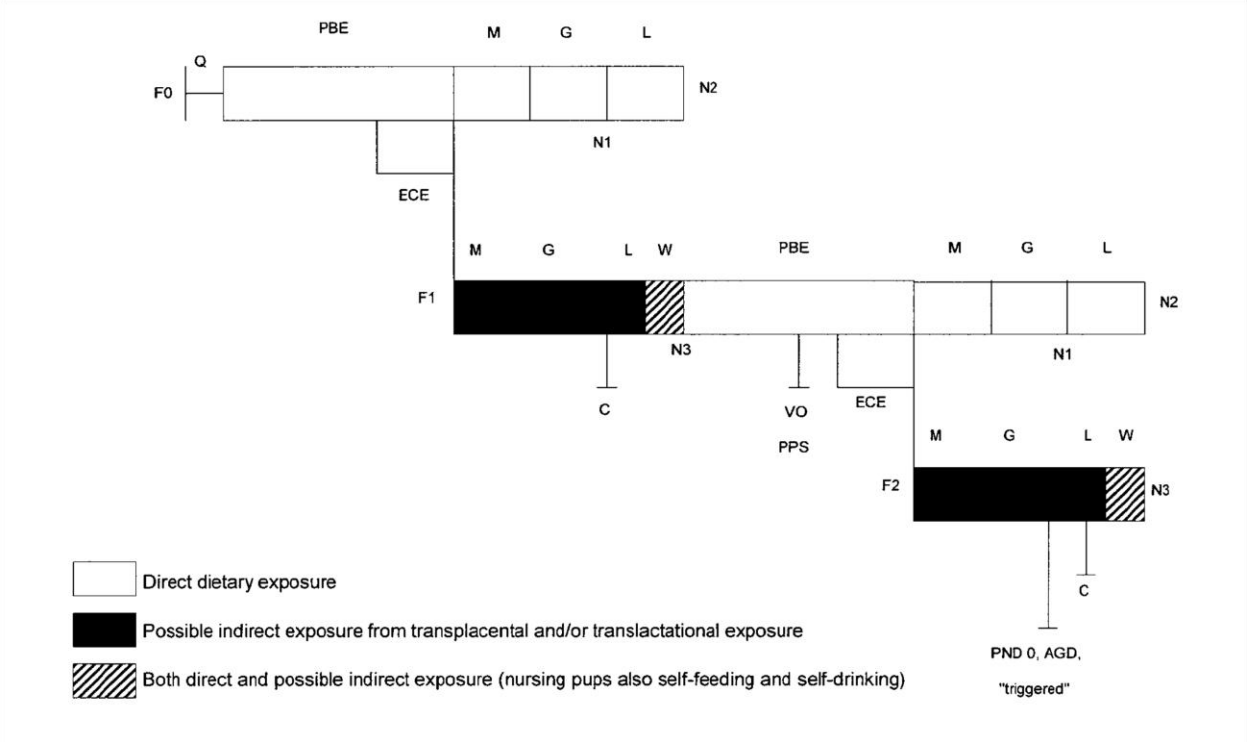


Figure 1. Diagrammatic representation of the current EPA test guideline for fertility and reproductive effects (OPPTS 870.3800).

Key: Q = quarantine; PBE = pre-breed exposure; ECE = estrous cycle evaluation; M = mating; G = gestation; L = lactation; W = weaning; N = necropsy; C = cull; VO = vaginal opening; PPS = preputial separation; PND = postnatal day; AGD = Anogenital distance.

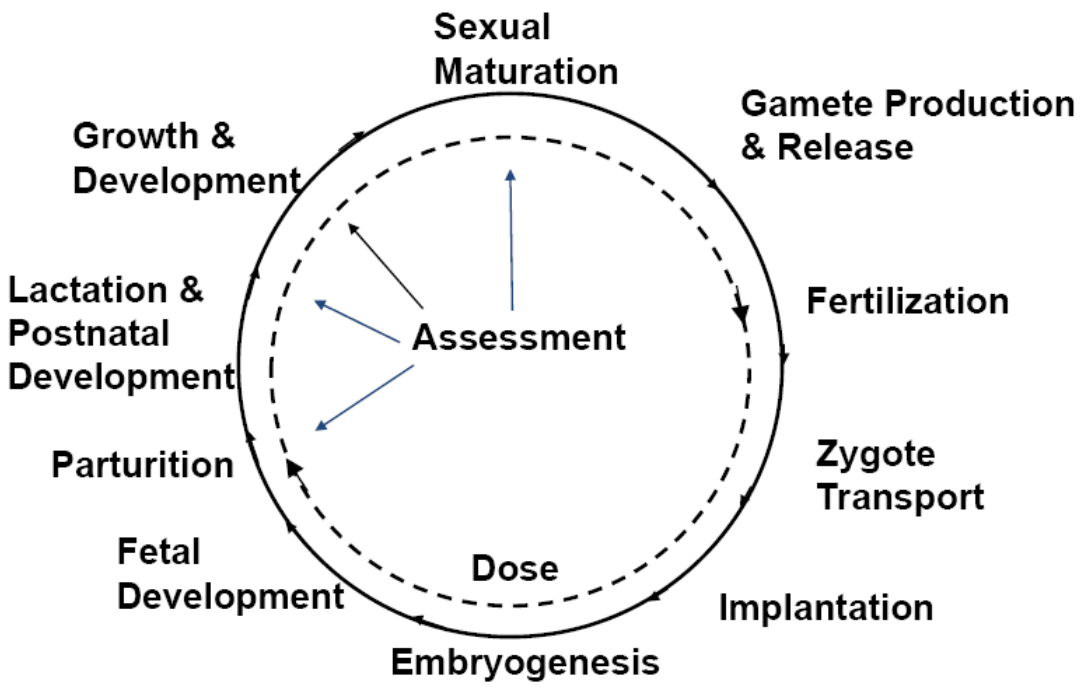


Figure 2. Diagrammatic representation of a typical multigeneration study with reference to the mammalian reproductive cycle. Dosing is continuous throughout the cycle with assessments made at multiple life stages.

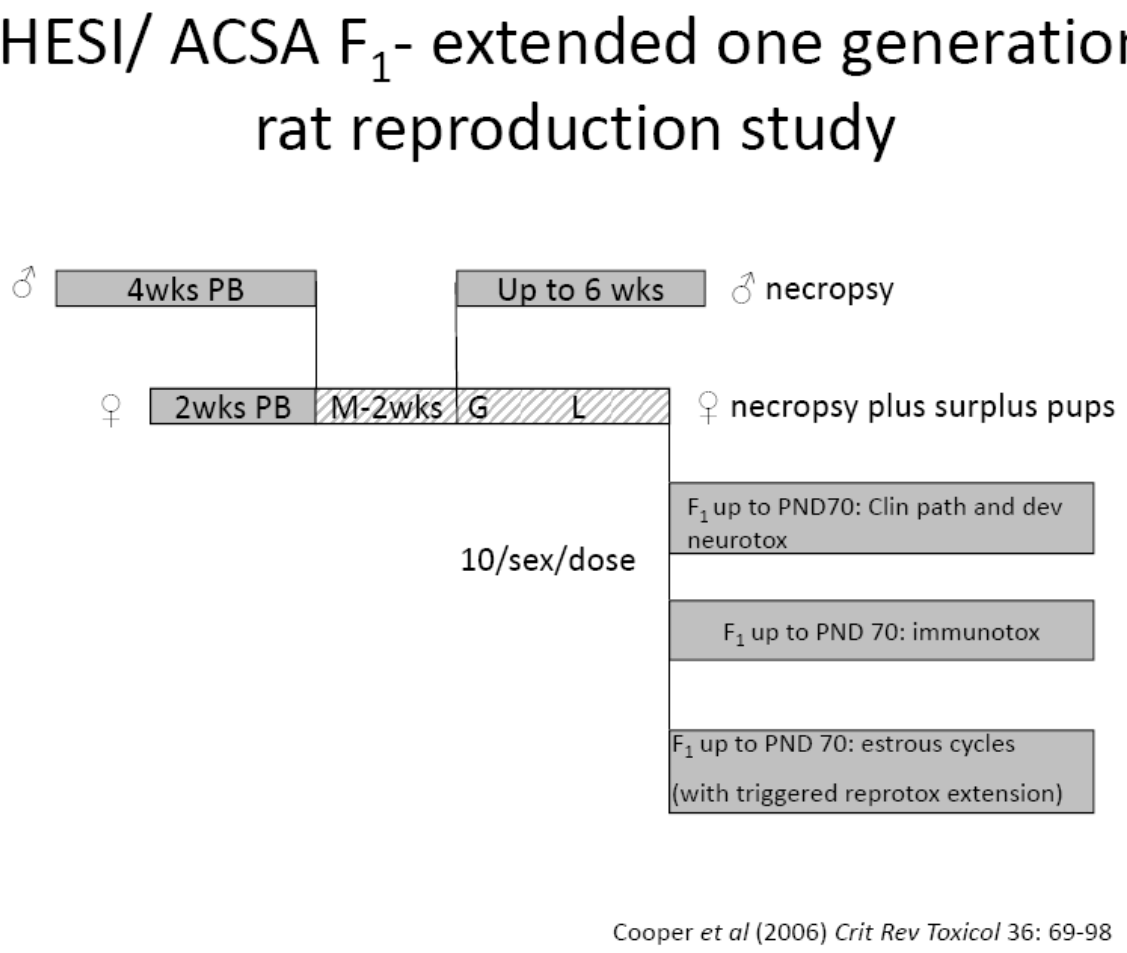
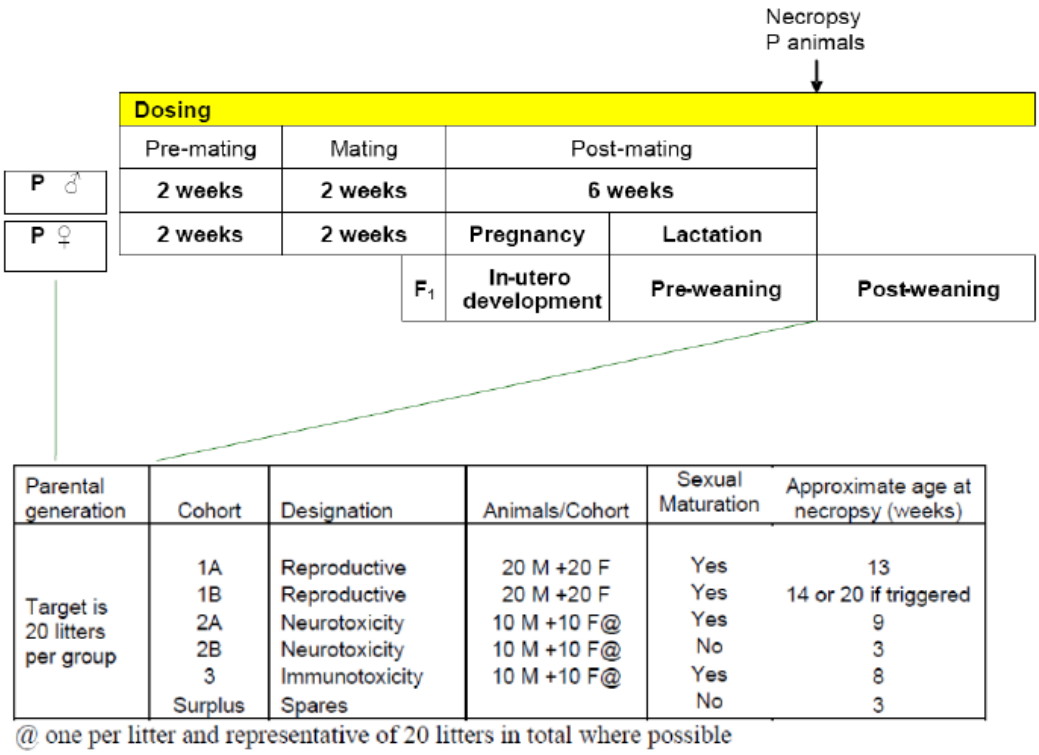


Figure 3. Diagrammatic representation of the Extended One generation reproduction study proposed by the Health and Environmental Science Institute for Agro-Chemical Safety Assessment.

Key: PB = pre-breed; M = mating; G = gestation; L = lactation; PND = postnatal day

OECD Proposed Design (version #28)



Parental generation	Cohort	Designation	Animals/Cohort	Sexual Maturation	Approximate age at necropsy (weeks)
Target is 20 litters per group	1A	Reproductive	20 M +20 F	Yes	13
	1B	Reproductive	20 M +20 F	Yes	14 or 20 if triggered
	2A	Neurotoxicity	10 M +10 F@	Yes	9
	2B	Neurotoxicity	10 M +10 F@	No	3
	3	Immunotoxicity	10 M +10 F@	Yes	8
	Surplus	Spares		No	3

@ one per litter and representative of 20 litters in total where possible

Figure 4. Diagrammatic representation of the Draft OECD Extended One Generation Reproduction Study.

Key: P = parental generation; M = males; F = females.

http://www.oecd.org/dataoecd/23/10/46466062.pdf Accessed Feb 16, 2011

NTP Modified One-generation Study

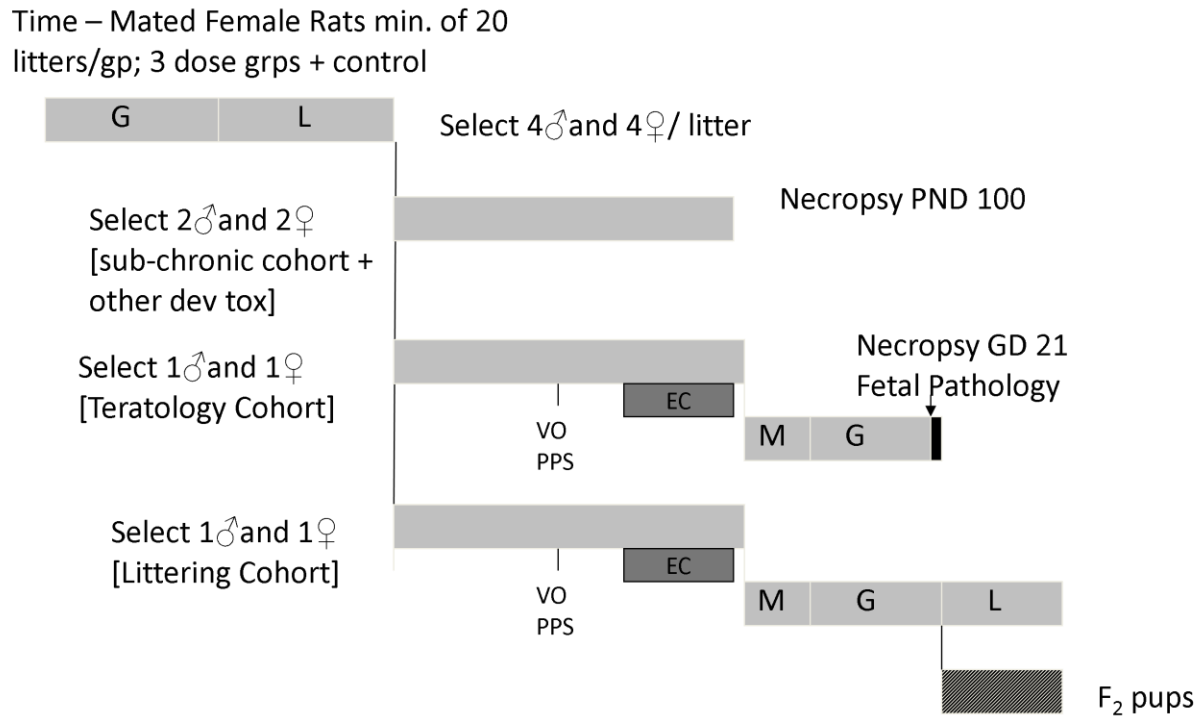


Figure 5. Diagrammatic representation of the Proposed NTP modified one generation reproduction study.

Only 10 pups per sex (on reaching adulthood) are required for the subchronic cohort and thus sufficient numbers of animals would be available for evaluations of other developmental toxicity that may include effects on the developing immune or nervous systems.

Key: G = gestation; L= lacation; PND = postnatal day; GD = gestation Day; M = mating; VO = vaginal opening; PPS = belanopreputial separation; EC = estrous cyclicity evaluation.